图书情报工作
第65卷 第3期 2021年2月

"一带一路"专题数据库建设调查与发展分析*

司莉 陈辰 周璟
武汉大学信息管理学院 武汉 430072

摘 要:[目的/意义]调查和分析"一带一路"专题数据库建设情况,为构建"一带一路"多语种、共享型数据库提供借鉴。[方法/过程]通过文献和网络调研法,选择30余种数据库作为分析对象,从资源组织体系、资源类型和内容、数据库建设方式以及多语言处理问题等维度进行系统分析,指出数据库建设具有资源组织方式不一、资源类型和内容多样、合作建设方式不同、页面语言较为单一的特点。[结果/结论]针对数据库建设存在的问题,提出采用多维度资源组织方式,注重资源内容的深度加工,加强国内外合作机制研究,及重视小语种资源采集和跨语种检索研究等方面启发性建议,为"一带一路"专题库建设提供参考。

关键词:"一带一路"数据库 建设方式 资源组织 合作方式 多语言资源
分类号:G250.74
DOI:10.13266/j.issn.0252-3116.2021.03.001

随着参与"一带一路"倡议的国家和地区数量的增多,"一带一路"网络不断被织密,"一带一路"专题数据库为"一带一路"信息流动和共享搭建"信息桥",在推动信息的互联互通方面发挥了重要作用。近些年,相关政府机构、教育和科研机构、信息服务机构及其他社会组织着眼于服务"一带一路"倡议目标,构建的"一带一路"专题数据库多达30余种。不过,"一带一路"专题数据库在多源异构与多语种资源融合方面还面临诸多挑战。目前"一带一路"信息化建设存在碎片化、低质化、信息不对称等问题,在满足用户需求方面还存在较大的改善空间[1]。数据库建设在数据源查找、数据质量保障以及多语言处理等方面还面临诸多难点[2]。语种资源较为单一,小语种资源相对欠缺[3]。虽然数据库内容侧重分明,但尚缺少有关"一带一路"沿线国家和地区深入、全面和系统的专题数据库[4]。

已有研究中尤其缺乏对数据库资源组织体系、资源类型和内容、建设合作方式等问题的系统分析,笔者以30余种"一带一路"数据库及服务平台为对象,从数据库的资源组织体系、资源类型和内容、建设方式和多语言处理问题等方面进行调查分析,以期更加全面地呈现"一带一路"专题数据库建设发展现状,针对存在的问题提出加强资源组织、内容建设、共建合作与跨语言检索的对策。

1 文献回顾

"一带一路"专题数据库研究主要集中在如下几方面:

(1)在数据库/平台系统架构方面,相关研究者对"一带一路"新型智库信息服务体系[5]、境外矿产资源勘查开发信息服务平台[6]、档案信息资源共享平台[7]、空间信息云服务平台[8]等的系统架构、技术路线和主要功能等问题进行了系统研究,突出了云计算、大数据库、移动应用、WebGIS等技术在"一带一路"数据库建设中的应用,这也是未来数据库技术应用的主流趋势。

(2)在数据库资源建设方面,相关研究者分别对"一带一路"新型智库[9]、丝路科技知识服务系统[4]、"一带一路"电影国别信息库[10]、"一带一路"留学生档案信息库[11]的资源采集、整合策略进行了研究,提出资源建设面临的问题集中在:资源采集标准的确定、资源整合深度不足、知识产权问题、小语种资源建设不足、资源利用率不高等方面,并根据问题提出了相应的

* 本文系教育部哲学社会科学研究重大课题攻关项目"'一带一路'沿线国家多语种、共享型经济管理数据库建设研究"(项目编号:19JZD020)研究成果之一。
作者简介:司莉(ORCID:0000-0003-1028-8338),教授,博士,博士生导师;陈辰(ORCID:0000-0002-0485-7525),博士研究生,通讯作者,E-mail:chenchensee@163.com;周璟(ORCID:0000-0002-6918-7163),硕士研究生。
收稿日期:2020-06-26 修回日期:2020-09-26 本文起止页码:4-12 本文责任编辑:徐健

司莉，陈辰，周璟.“一带一路”专题数据库建设调查与发展分析[J].图书情报工作,2021,65(3):4－12.

思考和建议。在多语种资源建设方面，主要集中在高校图书馆对小语种信息资源共建共享问题的探讨，如认为外语院校可依托联盟实施动态联合采购小语种资源，突出特色馆藏建设等方面加强小语种资源共建水平，通过制定合理经费预算和分配方案、打造多语种选书平台、采用“读者决策采购”的资源建设模式等方面加强小语种馆藏建设，提高资源共享水平[12－13]。

已有研究涉及“一带一路”专题数据库平台架构、资源建设、多语种等方面，为数据库建设提供了参考，但大多只将单一或者部分数据库作为分析样本，尚缺乏对已有数据库的全面、系统比较研究；且研究只涉及数据库建设的某方面，论述角度较为单一，内容较为分散，尚缺乏对资源组织体系、资源类型和内容、建设合作方式、多语种等问题的系统分析。

2 “一带一路”专题数据库建设的调查分析

笔者通过网络调研、文献调研和平台测试等方法，从资源组织方式、资源类型和内容、建设主体、建设方式以及语种方面对“一带一路”专题数据库建设现状进行调查分析。

“一带一路”专题数据库(或平台)，是相关机构为配合国家“一带一路”倡议，旨在收集和整合“一带一路”沿线国家和地区的政治、经济、文化、教育、科技等各领域信息资源而搭建的信息服务平台。通过查阅文献、相关报道及检索院校图书馆订购的数据库资源中初步获取“一带一路”相关数据库，通过访谈获取经管领域研究者的数据库使用情况，笔者确定了最终的调查对象。具体结果详如表1所示：

表1 “一带一路”专题数据库建设的总体调查

序号	数据库/平台名称	资源组织方式	资源类型	建设主体	建设方式	页面语种
1	中国一带一路网	综合型	新闻、政策文件、报告、数据	国家信息中心	开放合作、免费共享	中、英、俄、法、阿、西
2	一带一路生态环保大数据服务平台	综合型	资讯、政策法规、数据	中国生态环境部	合作共建(联盟合作)、免费共享	中、英
3	税收服务“一带一路”	资源类型－主题型	新闻、政策法规、视频、图片、报告	国家税务总局	合作共建(技术合作)、免费共享	中
4	上海市与“一带一路”国家经贸合作信息服务平台	综合型	新闻、政策法规、数据	上海市电子商务促进中心	独家建设、免费共享	中、英
5	湖北省一带一路公共服务平台	功能－主题型	新闻、报告	湖北省信息中心	合作共建(技术合作)、免费共享	中
6	“一带一路”研究与决策支撑平台	资源类型－主题型	新闻、政策法规、数据、报告	国务院发展研究中心信息网	合作共建(业务合作)、免费共享	中
7	“一带一路”物流综合服务平台	功能－主题型	资讯信息	郑州自贸片区管委会	合作建设(业务合作)、免费共享	中
8	一带一路冶金专题知识服务平台	资源类型－主题型	专利文献、数据、研究报告、新闻资讯	中国工程科技知识中心冶金分中心	合作共建(技术合作)、免费共享	中
9	“一带一路”共建国家标准信息平台	综合型	标准文献、新闻资讯	中国标准化研究院	开放共建、免费共享	中、英
10	“一带一路”大数据库“丝路信息网”	——	数据、论文、著作、研究报告、评论报道等	上海社会科学院与中国国际经济交流中心	合作共建(协议)、免费共享	中
11	丝路科技知识服务系统	资源类型－主题型	期刊、会议、图书、专利、报告、学位论文、视频等	西安交通大学	合作共建(联盟合作)、免费共享	中、英、俄、阿
12	北京大学“一带一路”数据分析平台	综合型	新闻、报告、会议、图表	北京大学	合作共建(项目合作)、免费共享	中
13	香港贸易发展局“一带一路”资讯网	功能－主题型	新闻资讯、政策法规、报告	香港贸易发展局与上海社会科学院合作	合作建设(协议)免费共享	中(简繁体)、英
14	“一带一路”国家基础设施发展指数信息服务平台	综合型	指数、报告、资讯、数据	中国对外承包工程商会	合作共建(技术合作)、会员制共享	中
15	一带一路资源中心数据库(BRIRS)	——	期刊论文和研究报告	美国EBSCO公司	合作共建(业务合作)、付费使用	英、日、韩、德等30种
16	一带一路专题数据库(BRIC)	——	电子图书	iG出版集团	合作共建(业务合作)、付费使用	英
17	一带一路中国文献整合平台	主题－主题型	学术期刊、博硕论文、会议论文、报纸、年鉴、工具书等	中国知网	合作共建(项目合作)、付费使用	中

（续表 1）

序号	数据库/平台名称	资源组织方式	资源类型	建设主体	建设方式	页面语种
18	一带一路统计数据库	区域 – 主题型	数据	中经网数据有限公司	独家建设付费使用	中
19	一带一路国际行业数据中心	区域 – 主题型	数据、报告	欧睿国际有限公司	独家建设付费使用	英
20	一带一路工业和信息化产业资源平台	综合型	资讯、政策法规、报告、数据（图表）、图书等	工信部电子工业出版社	独家建设免费共享	中
21	一带一路产业地图资源平台	资源类型 – 主题型	数据、研究报告、政策法规	工信部电子工业出版社	独家建设付费使用	中
22	新华丝路 – 一带一路国家级信息服务平台	综合型	新闻、报告、数据、期刊维普和万方会议、学位论文	新华社中国经济信息社	合作建设（备忘录）、部分付费使用	中、英、意、泰
23	新华丝路数据库	资源类型 – 主题型	新闻、数据、论文、报告、视频等	新华社中国经济信息社	独家建设付费使用	中、英
24	色诺芬（SINOFIN）一带一路专题库	综合型	数据	色诺芬信息技术有限责任公	独家建设付费使用	中
25	区域研究数据库 – 一带一路	资源类型 – 主题型	数据	北京金图公司	独家建设付费使用	中
26	清华控股“一带一路”	综合型	资讯、图片	清华控股集团	独家建设免费共享	中
27	Wind 一带一路数据库	资源类型 – 主题型	数据	万德信息技术股份有限公司	独家建设付费使用	中
28	RESSET 一带一路信息平台	资源类型 – 主题型	新闻、政策法规、报告和数据	锐思数据科技有限公司	独家建设部分付费使用	中
29	RESSET 一带一路数据库系列	区域 – 主题型	数据	锐思数据科技有限公司	独家建设付费使用	中（简繁体）、英
30	CSMAR“一带一路”研究数据库	主题 – 主题型	数据	国泰安信息技术有限公司	合作共建（合作数据）、付费使用	中
31	CNRDS 一带一路研究数据库 – BRRD	区域 – 主题型	数据 + 部分论文摘要	上海经禾信息技术有限公司	合作共建（合作数据）、付费使用	中
32	“一带一路”数据库 – 国家前沿战略支撑平台	综合型	图书、新闻、数据（图片）	社会科学文献出版社	合作共建（业务合作）、部分付费使用	中（简繁体）
33	“一带一路”经济信息共享平台（BRInfo）	——	新闻资讯	新华丝路负责创建和运营	开放合作（倡议合作）、免费共享	英

注：“——”表示因数据库无法访问或获取说明资料而无法确定资源组织方式

当前“一带一路”数据库建设已初具规模，且建设主体多样、领域涉及广泛、资源类型多样，可从多方位、多角度满足不同领域、不同用户群体的信息需求，数据库具体建设情况分析如下：

2.1 关于数据库的资源组织方式

通过直接访问数据库和间接依据数据库说明等方式，分析资源目录展开一二级层次的先后顺序，29 个“一带一路”数据库（4 个数据库无法访问或获取说明资料而排除在外）的资源组织方式可总结为以下 5 种：

2.1.1 资源类型 – 主题型

首先按资源类型将信息划分若干子库或一级栏目，再根据专题、领域、区域、国家等主题进行资源的细化组织。采用该方式的数据库共 9 个，占总样本的 31.0%，包括新华丝路数据库、区域研究数据库、丝路科技知识服务系统等。如新华丝路一级类目按智讯、咨询、资讯、视讯、情报、案例、论文、报告、项目、数据和园区等资源类型进行设置，各个栏目再按不同的主题类别进行细分，如资讯按“政治外交”“营商环境”“政策法规”等主题进行细化组织。

2.1.2 区域 – 主题型

首先按“一带一路”沿线国家或地区所属区域将其划分为若干子库或一级类目，再按主题对区域资源进行细化组织，采用的数据库共 4 个，占总样本的 13.7%，包括一带一路统计数据库、RESSET 一带一路数据库系列等。如一带一路统计数据库首先划分新亚欧大陆桥、中蒙俄、中国 – 中亚 – 西亚、中国 – 中南半岛、中巴、孟中印缅六大经济走廊和 21 世纪海上丝绸之路共七大线路，将沿线国家和省市所属线路划分子库，再围绕“一带一路”五大合作重点（经济概况、能源环境、设施联通、贸易畅通和民心相通）组织指标结构，

结构清晰，并支持逐层浏览和查询[14]。

2.1.3 功能－主题型

首先从满足用户需求出发设置服务栏目，再根据主题对满足特定功能的资源进行组织，采用这种方式的数据库共3个，占总样本的10.3%，包括香港贸易发展局的一带一路资讯网、湖北省一带一路公共信息服务平台等。如湖北省一带一路公共信息服务平台，搭建了“信息服务”“资讯服务”“政企服务”和“互动服务”四大服务板块，各板块再按相应的主题进行资源细分。此类平台相对于内容，更强调功能性和服务性。

2.1.4 主题－主题型

资源的一级类目或二级以上类目都按主题进行划分，采用此体系的数据库共2个，占总样本的7%，包括中国知网的一带一路中国文献整合平台和国泰安的CSMAR“一带一路”研究数据库。如CSMAR“一带一路”研究数据库设置“沿线国家社会经济发展指标”和“一带一路特色指标”两个一级主题类目，其类目下又有“人口”“国民经济核算”“就业和投资”等细化指标。

2.1.5 多维度综合型

多维度综合型是指资源组织没有单一的标准，而采用多维度方法对资源进行归集，采用这种方式的数据库共11个，占总样本的38.0%，包括国家信息中心的中国一带一路网、社会科学文献出版社的“一带一路”数据库、新华丝路信息服务平台等。如社会科学文献出版社设置包括资讯库、理论库、国家库、省域库、战略库、专题库、指数库、史话库和专家库8个子库，其子库划分并没有依据上述标准中的某一类进行，而是将区域国家、资源类型、主题等交叉组织其类目资源。

已建数据库的资源组织方式不一，“多维度综合型”是使用最多的资源组织方式，该方式类目设置灵活、资源的可扩展性较强，多适合大型平台对多区域、多领域、多类型的资源的组织。“资源类型－主题型”的使用量排名第二，该方式类目设置清晰，方便用户按数据、新闻、报告等不同类型资源进行浏览和检索，多适用于具有多种类型资源的数据库。“区域－主题型”的使用量排名第三，该方式结构清晰，查询方便快捷，多适用于只包括统计数据的数据库。

2.2 关于数据库的资源类型与内容

2.2.1 资源类型

从整体上看，“一带一路”数据库资源类型包括新闻资讯、政策法规、电子图书、期刊论文、学位论文、会议论文、统计数据、研究报告、标准文献和专利文献等多种类型。不同建设主体侧重的资源类型具有明显差异。学术教育和政府机构构建的数据库，侧重采集新闻资讯、政策法规、研究报告、标准文献等类型，此类信息更新较快，时效性较强，且开放共享程度较高，多数可供用户免费使用；信息服务提供商构建的数据库，侧重采集数据、论文、电子图书和专利文献等学术资源，更新相对较慢，学术性较强，但受限程度较高，一般需要授权才能访问。

2.2.2 资源内容

结合数据库访问和依据数据库说明资料等方式，对数据库的资源内容体系进行调查，按13个分类框架对资源内容进行统计。共调查了26个数据库（有7个数据库因无法访问、无法获取说明资料、有自身特有分类体系而不适用此框架的排除在外），统计过程将反映同一内容的栏目进行归一，如将专家观点、专家视点、记者观察、专家库、专家资源库、研究人员档案库等归为专家资源，将实践探索、实务、案例等归为案例信息。资源内容分布如图1所示：

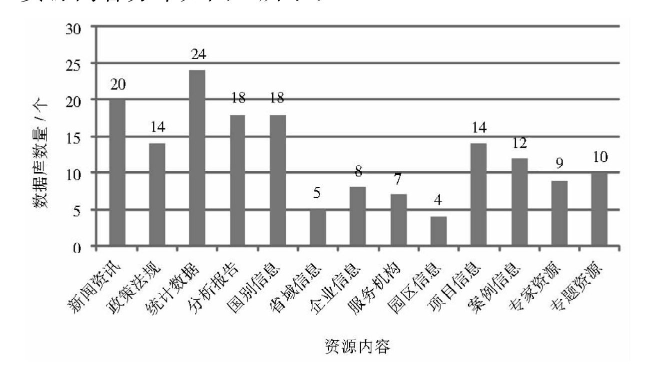


图1 数据库中的资源内容统计数量分布

不同的资源内容在数据库中的出现频次并不一致，其中“统计数据”的频次最高，出现频率为24次，占92%，是各类平台广泛关注的资源内容。排2－4位的依次分别为“新闻资讯”（20次，占77%）、“国别信息”（18次，占69%）和“分析报告”（18次，占69%），这类信息大多是了解“一带一路”倡议内涵及规划布局所涉及的背景信息，如涉及国家区域的经济、社会、文化发展数据，商务部、外交部发布的风险评估报告等，而国内企业抱团出海、集群式“走出去”所需要的“企业信息”（占31%）、“服务机构”（占27%）和“园区信息”（仅占15%）则相对匮乏。此外，部分数据库提供“专题资源”（占38%），主要为各个平台特有的“专题栏目”，如特色专题库、特色成果库和热点专题等。

2.2.3 统计数据分布

从“统计数据”出现频率可以看出，各类数据库都注重采集的资源类型，在政府决策、科学研究、企业投

资过程具有举足轻重的作用。笔者根据国家统计局的22个数据分类指标,对20个数据库中(共24个数据库包括统计数据,4个数据库因是图表数据、资讯报道数据、标准数据、指数数据,而不适用此指标进行统计,固排除在外)统计数据进行指标分布统计,结果如图2所示:

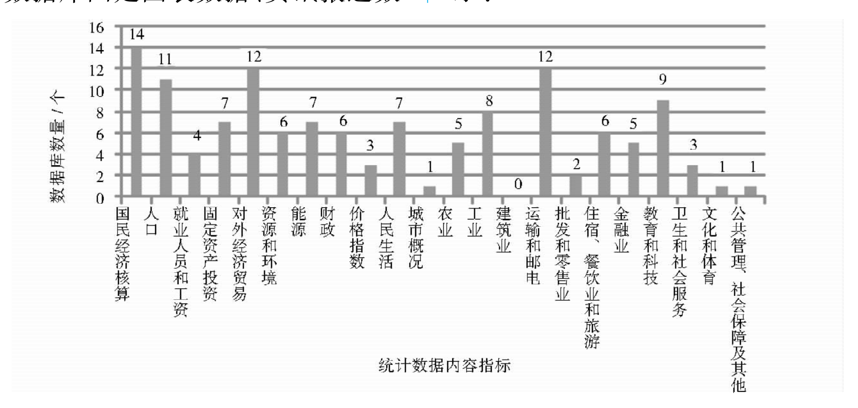


图2 统计数据指标在数据库中的数量分布

从图中可以看出,统计数据指标分布并不平衡。首先出现频次较多的为"国民经济核算"数据,主要是涉及国民生产总值(GDP)、三次产业、居民消费水平以及国际收支情况方面的数据;其次是"对外经济贸易"和"运输和邮电"数据,出现频次位列第二;再者,"人口"数据位列第三。除了上述出现频次较高的指标之外,涉及城市概况、文化和体育、公共管理、社会保障及其他则很少涉及,仅有1个数据库,还没有数据库采集建筑业的数据。

此外,还出现超出国家统计局22个指标以外的数据,可视为"一带一路"特色发展指标,如官方发展援助、外交、公共部门(战争、军费、武器等信息)、扶贫和犯罪统计等,提供"一带一路"沿线国家和地区有关环境评估和风险预测等方面的背景数据。

2.3 关于数据库的建设方式

2.3.1 数据库建设或依托主体

数据库建设或依托主体涉及信息服务提供商、教育机构、学术机构、社会组织和政府机构等多种类型。

(1)信息服务提供商。主要以营利为目的的数据库商和出版商为主,其构建的"一带一路"专题数据库,一般实施付费模式,受限程度较高。此类主体构建的数据库共19个,占总样本的57.6%,如美国EBSCO公司、iG出版集团、欧睿国际有限公司、社会科学文献社、北京金图公司、锐思数据科技有限公司和中国知网等。

(2)相关政府部门。这些部门积极响应国家"一带一路"倡议,为企业、机构社团和个人等积极参与"一带一路"建设提供信息服务和互动交流平台。此类主体构建的数据库为7个,占总样本的21.2%,如国家信息中心、国家税务总局、中国生态环境部等国家机构,湖北省信息中心、上海市电子商务促进中心等地方机构。

(3)教育、学术和科研机构。利用专业研究及团队合作优势,汇集整合外部资源,开展数据相关研究,并形成分析报告,此类主体构建的数据库为5个,占总样本的15.2%,如北京大学、西安交通大学、上海社会科学院、中国标准化研究院等机构。

(4)社会组织或行业组织。基于特定目的构建的服务于政府和企业的数据库,此类主体构建的数据库为2个,占总样本的6%,如香港贸易发展局和中国对外承包工程商会,数据库访问一般采取会员机制,不同的会员级别具有不同的数据访问权限。

可见,约有半数的数据库属于信息服务商构建的(57.6%);其次为相关政府机构搭建的共享型平台(21.2%);再次为教育、学术和科研机构构建的数据库(15.2%),少量的社会组织或行业组织为满足特定行业和领域的信息需求而构建的数据库(6%)。

2.3.2 数据库建设方式

数据库建设方式可归纳为3种类型,分别是独家建设与运营、合作建设运营和开放式建设运营。

(1)独家建设与运营,通常由一家机构独立完成数据库的建设和运营工作,采用此类方式的数据库有12个,约占总样本的36.4%,其中大多数为数据库商,凭借多年的资源累积和数据库建设专业经验,独立完成"一带一路"数据库的建设。

(2)合作建设和运营,由两个或两个以上机构联合构建和运营数据库,采用此类方式的数据库有18个,约占总样本的54.6%。通过对其合作方式进行调研,发现如下几种类型:

第一类型——业务合作方式:基于数据库建设主体已有的业务合作网络,进行“一带一路”数据库建设。采用此方式构建的数据库为 5 个,占比 15.1%。如上海社会科学院已与全球 90 多家智库和学术机构建立了合作关系网络,其中有近 30 家位处“一带一路”沿线国家,为获取“一带一路”沿线国家信息进行“丝路信息网”构建奠定了合作基础[15]。

第二类型——技术合作方式:由技术力量较为薄弱的主体与商业技术公司进行合作开发,采取此方式构建的数据库有 6 个,约占总样本的 18.1%。如湖北省一带一路公共服务平台由武汉世纪金桥安全技术有限公司提供技术支持。此外,国泰安信息技术有限公司也对外提供技术平台,与专业领域的数据库公司、相关领域的专家和学者进行数据合作,支持“一带一路”研究基础数据库的构建。

第三类型——项目合作方式:基于项目合作关系以及项目阶段成果,进行“一带一路”数据库建设。采取此方式构建的数据库有 2 个,约占总样本的 6%。如 2014 年中国知网与国家广电总局联合运作“丝路书香工程”－“中英文丝路文献数据库多国合作项目”,该项目旨在联合丝路相关国家共同开发出版中英文丝路文献数据库,2018 年中国知网整合了项目的先期阶段性成果,构建了一带一路中国文献整合平台[16]。

第四类型——协议/备忘录合作方式:通过两家或多家机构签订双边或多边合作协议,达成数据库共建共享合作意愿。采取此方式构建的数据库有 3 个,约占总样本的 9%。如中国经济信息社与泰国商业信息平台－Business Today,在泰国曼谷签署合作备忘录,借此,新华丝路信息服务平台的中国经济信息及“一带一路”建设相关信息以泰语形式呈现在 Business Today 平台上[17]。

第五类型——联盟合作方式:通过联盟约束机制,促进合作伙伴间在数据库共建共享方面的共识、合作和一致行动,采取此方式构建的数据库有 2 个,约占总样本的 6%。如中国生态环境部着手建设生态环保大数据服务平台和“一带一路”绿色发展国际联盟,通过联盟合作伙伴关系促进国际平台的共同搭建和维护,从而实现信息的共享[18]。

(3)开放式建设和运营,秉承开放和包容理念,多机构合作参与数据库的建设。采用此类建设方式的数据库有 3 个,约占总样本的 9%。如国家信息中心明确指出,“充分秉持开放合作的思想,欢迎各地方、沿线国家共同参与官网建设”[19]。由新华社牵头发起和推动的“一带一路”经济信息共享网络,在成员之间实行经济信息内容共建和免费共享机制,目前已吸引来自 26 个国家的 30 多家机构参与经济信息共享平台的共建和共享工作[20]。

在 3 种数据库建设方式中,采用合作方式建设的数据库数量最多,其次为独家建设的数据库,这两种是较为常见的方式,但合作范围和参与主体有限,且多限于国内机构的合作,国外合作力度不足。开放式建设和运营方式强调跨国家、跨区域、跨领域的多方主体共同参与,但目前实施程度有限。

2.4 关于数据库的多语言问题

“一带一路”沿线国家官方语言多达 50 余种,且不少是小语种,语言问题成为数据采集、组织、挖掘和利用的最大障碍。大多数据库通过设置多语言页面版本来解决语言障碍问题。图 3 为数据库页面所用语种的分布。

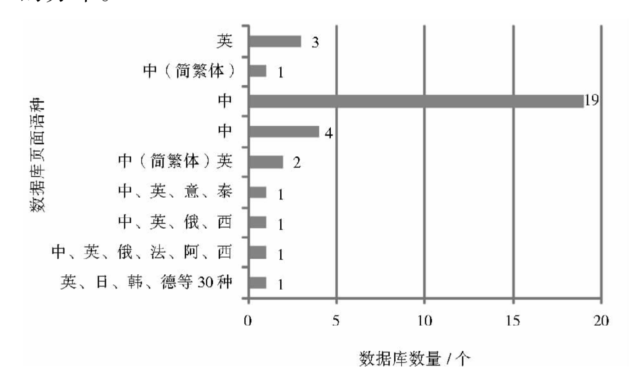


图 3 “一带一路”数据库页面所用语种的分布

由图 3 可见,支持多语言页面显示的数据库数量较少,仅有“中国一带一路网”正式发布了联合国六大官方语言中文、英语、俄语、法语、西班牙语、阿拉伯语页面。“丝路科技知识服务系统”,支持中、英、俄、西 4 种语言页面,美国 EBSCO 公司的数据库支持英、日、韩、德语等 30 种语言,新华丝路－一带一路国家级信息服务平台支持中、英、意大利语和泰语页面。支持中英文语言页面的数据库共 6 个,只支持中文页面的数据库数量最多,包括中文与中文(简繁体)页面的共 20 个。

对于不同语言版本的多语种资源的页面揭示看,只对资源的题录元数据进行了语种转换,如以丝路科技知识服务系统俄语页面为例,将英文资源的题录元数据翻译成俄语,如“作者”“来源”和“主题词”等转换为“автор”“источник”和“Тематический термин”,但是字段值及原文仍然是原有的语言形式。

从跨语言检索功能实现上看,尚未发现支持此功能的“一带一路”数据库。虽然 EBSCO“一带一路”专

题库的期刊论文题目、关键词和作者等通常包含多语种的翻译,如中文和英文题录、西班牙语和英文题录,但只能够实现题录层面的跨语言检索。

3 "一带一路"专题数据库建设存在的问题及建议

笔者通过对"一带一路"数据库建设发展情况的系统梳理,借鉴已有数据库在资源建设合作等方面的有益经验和方法,识别目前数据库建设存在的不足与薄弱环节,并针对问题提出相应的改进对策,为构建"一带一路"多语种、共享型数据库提供参考和启示。

3.1 资源组织方式不一,宜采用多维度资源组织方式

已建数据库的资源组织方式不一,不同的数据库根据功能定位和用户需求采用了不同的资源组织方式。"多维度综合型"是目前使用最多的资源组织方式,该方式类目设置灵活、资源的可扩展性较强,适合大型平台对多区域、多领域、多类型的资源的组织策略。"一带一路"数据库涉及的领域范围广、资源类型多样,且需要同时满足政府、企业、科研用户等的不同群体的信息需求,因此宜采用多维度资源组织方式,对内容进行多角度、多层次揭示,方便用户根据个人意向和偏好选择搜索途径,从而高效快捷地找到所需资源。

首先,可借鉴多重列类的方式,对资源进行多维度揭示,为用户提供多种检索途径。如"一带一路"研究与决策支撑平台设置了国别、省市、行业、案例、项目、法规、资讯等平行类目,同时又分别设置"国别"和"省市"分类标签对全库内容进行自动聚合分类,用户通过标签就可以浏览相关国家或省市的全部案例、项目、资讯等信息。因此,在广泛建设信息站点的同时,也要重视多重列类,多视角、多层次、全方位对数据库资源进行立体梳理和分类呈现。

其次,基于可扩展性原则设置分级栏目体系,循序渐进、主次分明的推进资源建设。如国研网"一带一路"研究与决策支撑平台,同时设置主要栏目和辅助栏目,前者包括报告、案例、行业、法规、项目、数据以及图表,是资源组织的重点和核心,而后者包括资讯、规划方案、机构、大事记和重点报告推荐,是资源组织的补充和扩展,这样既可以分期推进资源建设,又可随时在辅助栏目增添类目,为资源组织的可扩展性预留空间。

3.2 资源开发力度不足有待深化,注重资源内容的深度开发

已建部分数据库存在资源同质化采集以及内容开发力度不足等问题。当前有关"国家背景""政策法规""新闻资讯"等内容重复建设问题很突出,而企业需要的投资信息、园区信息等则相对匮乏。已有数据库对于国家和省市信息的采集,多为复制粘贴商务部、外交部网站的信息,有关的新闻资讯等也缺少相应的主题分类和组织,给用户信息检索带来较大不便。因此,宜针对"一带一路"特点对数据进行深度加工处理,形成有影响力的信息产品,满足用户个性化检索需求。

首先,加强国家机构掌握的基础性信息的开放,通过链接等方式实现对这些信息的关联,避免重复的复制粘贴工作,将更多精力放在资源的加工和产品开发上,形成具有核心竞争力的"一带一路"信息产品。如上海社会科学院与中国国际经济交流中心构建的"丝路信息网",既重视对基础数据的采集工作,同时又注重数据的分析和加工工作,研制数据分析动态报告等智库产品,为政府、企业、学术机构等提供形势判断和决策咨询服务。

其次,加强数据库之间相同主题不同平台、不同类型间的资源关联和整合,满足用户知识检索和发现的需求。"一带一路"的专题数据库很多,但相同主题信息在各个数据库中呈碎片化分布,各平台间缺少互操作和关联,难以满足用户"一站式"检索需求。可借鉴新华丝路"去新闻化"的资源整合模式,通过信息链的方式将同一主题的不同资源进行串联,形成更深入的信息主题阵列,满足用户的差异化、多元化信息需求[21]。

再者,运用人工智能、大数据挖掘等技术提高海量资源的组织效率。如"一带一路经济共享平台"通过人工智能技术自动对海量新闻资源进行元数据的分类抽取,解析标题、摘要、关键词、作者、类别等字段,以供用户检索和浏览。

3.3 资源建设合作力度有待增强,加强国内外合作机制研究

部分数据库采用多种合作方式推进数据库的建设,但合作范围和参与主体有限,多限于国内相关机构,国外合作十分有限。对于倡导跨国家、跨区域、跨领域的多方主体共同参与的开放式建设和运营方式,目前实施程度有限,限制了资源整合和共享的范围。日后宜加强国外共建共享合作机制研究,形成内外联动的数据库建设合作体系,才能让信息共享赋能各方优势互补和惠益共享。

首先,依托国际组织和机构,协调不同国家和地区

的数据合作。鉴于亚投行、上海合作组织、金砖银行等国际组织的成员国与中国具有良好的合作基础，在推动信息交流和合作方面具有得天独厚的优势。推动这些国际组织建立专门机构，支持和协调不同国家的信息共享合作，为促进跨国信息交流和整合提供通道和便利。

其次，加强国际间的合作机制研究，增强数据库共建共享的动力。如"一带一路"经济信息共享网络倡议实行内容的共建和资源的免费共享机制，共享网络成员有义务向平台提供可发布的经济信息，同时享有使用平台内容的权益，这种义务和权益平衡机制对于激发成员机构进行信息共建共享具有推动作用。此外，"数据银行"模式创造了一种类似共享经济和电子商务的机制，将零散、碎片化资源在一个平台进行呈现、匹配和交易，优化了资源供需匹配效果，利于资源的开放式整合[22]。

3.4 多语种资源相对匮乏，重视小语种资源采集和跨语言检索

已有数据库中的小语种资源收集相对匮乏，仅有少量国际出版商涵盖了多种小语种资源，但远不能满足国内外用户的多语种信息需求。已有数据库支持多语言页面显示的数据库数量较少，多为只支持中文页面的数据库。从对不同语言版本的多语种资源的页面揭示看，只对资源的题录元数据进行了语种转换，而尚未实现资源的跨语言翻译和展示。从跨语言检索功能实现上看，尚未发现支持此功能的"一带一路"数据库，只有部分支持题录的跨语言检索。

首先，应加强对小语种国家资源的识别和采集。一方面，可通过翻译机器或软件对小语种资源进行翻译识别，再结合人工进行主题取舍和质量甄别，从而实现大范围多语言信息资源的内容聚合。另一方面，可通过与小语种国家的相关机构和人员构建合作关系，多渠道、多途径进行信息征集。

其次，进一步加强对跨语种信息检索技术、跨语种智能翻译技术、经济管理语料库建设等方面的系统调研，结合图书情报跨语言学术信息检索软件的技术积累，整合多个学科领域专家学者进行系统而专深的联合攻关研究，从而实现"一带一路"数据库跨语种的智能翻译和检索功能，突破用户面临的语言障碍。

4 结语

已有"一带一路"专题库中，不乏可以提供示范应用的典型平台，为专题数据库建设提供了宝贵经验，但在满足用户多语种、集成化信息需求等方面还存在较大的改善空间。通过对已有数据库的资源组织方式、资源类型和内容、建设合作方式以及多语言问题等方面进行系统梳理，总结了数据库建设特点，并识别和发现数据库建设存在的不足，并据此提出采用多维度资源组织方式，注重资源内容的深度开发，加强国内外合作机制研究，及重视小语种资源采集和跨语种检索等方面的研究建议，从而为建设和开发"一带一路"专题数据库提供参考和启示。

参考文献：

[1] 张展."一带一路"信息服务平台建设探析[J].辽宁经济，2015(7):26-27.
[2] 于施洋，杨道玲，王璟璇，等."一带一路"数据资源归集体系建设[J].电子政务,2017(1):8-14.
[3] 严丹，马吟雪."一带一路"专题数据库的建设现状及开发策略研究[J].图书馆学研究，2017(12):40-47.
[4] 李娟，张雪蕾，陈楠楠，等. 丝路科技知识服务系统"一带一路"专题数据库的构建与实践[J]. 数字图书馆论坛，2019(2):37-42.
[5] 赵益维，赵豪迈. 大数据背景下"一带一路"新型智库信息服务体系研究[J].电子政务,2017(11):72-80.
[6] 陆婧，董少春，王丽娟，等. "境外矿产资源勘查开发信息服务平台"的设计与实现[J].实验室研究与探索，2019(3):222-226.
[7] 卞咸杰.大数据时代档案信息资源共享平台构建的目标定位与实现[J].档案管理，2020(5):75-76.
[8] 胡金晖，秦智超，石磊，等. 空间信息云服务平台架构及应用研究[J].中国电子科学研究院学报，2016(1):51-58,83.
[9] 赵豪迈，赵益维. "一带一路"新型智库信息资源建设研究[J].情报探索，2020(2):63-66.
[10] 吕思诗."一带一路"视域下电影产业信息服务研究——以上海图书馆"一带一路电影国别信息库"为例[J].图书馆研究与工作,2020(2):55-59.
[11] 郑慧，黄艳芝."一带一路"倡议下高校留学生特色档案信息资源数据库建设研究[J].档案管理，2020(2):75-77.
[12] 王明惠，刘延莉. "一带一路"背景下外语院校图书馆非通用语种文献资源建设的困境与对策[J].晋图学刊，2017(2):18-23.
[13] 杨晓景. 高校图书馆收藏中国周边国家语种文献现状的调查研究[J]. 大学图书馆学报,2014(1): 54-58,70.
[14] 中国经济信息网. 中经网一带一路统计数据库使用手册[EB/OL].[2020-05-07]. https://ydyl. cei. cn/xml/introduction. pdf.
[15] 上海社科院:倾注智力丝路建设[EB/OL].[2020-05-07]. http://www. shekebao. com. cn/shekebao/n440/n452/u1ai12184. html.
[16] 清华控股. 同方知网"中英文丝路文献数据库多国合作项目"[EB/OL]. [2020-04-07]. http://zt. thholding. com. cn/in-

dex/show/contentid/29. html.
[17] 新华丝路. 中经社携手泰国商业信息平台助推“一带一路”经济信息在泰传播[EB/OL]. [2020 - 04 - 07]. https://www.im-silkroad.com/news/p/388648.html.
[18] 人民网. 环保部:生态环保大数据服务平台建设启动,助力“一带一路”国家绿色发展[EB/OL]. [2020 - 03 - 07]. http://env.people.com.cn/n1/2017/0720/c1010-29418704.html.
[19] 中国一带一路网. 中国推进“一带一路”建设官方网站:中国一带一路网[EB/OL]. [2020 - 03 - 08]. https://www.yidaiyilu.gov.cn/sy/gxwm/5059.htm.
[20] 新华社. 打通经济信息通道 共享“一带一路”机遇——写在“一带一路”经济信息共享网络成立之际[EB/OL]. [2020 - 04 - 07]. https://baijiahao.baidu.com/s? id = 1637456210870345414&wfr = spider&for = pc.
[21] 任玮玮. 新华丝路综合信息平台的实践探索[J]. 中国记者, 2019(5):32 - 34.
[22] 丁波涛. 基于数据银行的“一带一路”信息资源整合研究[J]. 情报理论与实践, 2018(12): 88 - 92.

作者贡献说明:

司莉:负责论文主题的选取和指导,修改论文;
陈辰:负责资料的整理和分析,撰写论文;
周璟:负责资料的收集。

Investigation and Development Analysis of “the Belt and Road” Thematic Database Construction

Si Li　Chen Chen　Zhou Jing
School of Information Management, Wuhan University, Wuhan 430072

Abstract: [**Purpose/significance**] Investigating and analyzing the construction of “the belt and road initiative” thematic database to provide reference for the construction of “the belt and road initiative” multilingual and shared database. [**Method/process**] Through literature and network research, more than 30 kinds of databases were selected as analysis objects, and systematic analysis was carried out from the aspects of resource organization systems, resource types and content, database construction methods, and multilingual processing problems. The construction of database is characterized by different ways of information organization, diverse types and contents of resources, different ways of cooperative construction, and single page language. [**Result/conclusion**] The paper proposes adopting multi-dimensional resource organization methods, in-depth processing of resource content, strengthening research on cooperation mechanisms at home and abroad, and attaching great importance on the collection of small-language resources and cross-lingual retrieval research.

Keywords: belt and road databases　mode of construction　resource organization　mode of cooperation　multilingual resources